

# レビューには何が書かれていて、読み手は何を読んでいるのか？

安藤 まや(慶応義塾大学 政策・メディア研究科) 関根 聡(ニューヨーク大学)

maya@sfc.keio.ac.jp, sekine@cs.nyu.edu

## 1. はじめに

インターネット上にある商品のレビューが購買の意思決定に影響を与えることが注目されている。レビューを使用した従来の研究では、レビューに含まれる単語や N-gram などのシンプルな素性を使用した、ポジティブ / ネガティブの自動分類、スパムか否かの判定、商品の属性や書き手である評価者の推定といったレビューの書き手の視点からの側面が研究対象となってきた。しかし、レビューが新規顧客に与える影響を考慮すると、書き手が書いた文章の特徴を把握することに加え、読み手の受ける印象という視点から研究することも重要である。読み手の視点からレビューを丁寧に分析すると、レビューには商品の感想のみならず、買い手の背景、商品の推薦などの様々な内容が含まれていることがわかる。従って、まずはレビューに含まれる文のタイプを把握し、それぞれのタイプの文に対する読み手の印象、特に購買意欲をかき立てるようなインパクトのある文の分布を把握する必要がある。しかし、このような視点からのレビューの分析は著者の知る限り行われてきていない。また、上記の分析結果から読み手にインパクトを与える文には、その商品の特殊性、描写の具体性、書き手の満足度合いが影響していることがわかった。これらの点を定量的に明らかにするために、本論文では、以下の3点についての分析を行う。

1. そもそもレビューには何が書かれているのか？
2. どのようなタイプの文に対して、読み手は購買意欲をかき立てられるほどのインパクトを受けるのか？
3. 商品の特殊性、描写の具体性、書き手の満足度合いなどのくらい読み手のインパクトと相関があるのか？

本研究では、楽天(株)が公開している楽天市場の 15 カテゴリーに含まれる商品と楽天トラベルのレビューの内容を分析し、商品の種類に左右されない共通の分類を 23 項目作成した(3 節)。そして、楽天トラベルのレビューに対し、この分類を基に読み手が受けるインパクトとの関係を調査した(4 節)。また、楽天トラベルのレビューの複数のトピックにおいて、商品の特殊性、描写の具体性、書き手の満足度合いと読み手が受けるインパクトの相関性を調査した(5 節)。

## 2. 関連研究

評判分析ではレビューをポジティブ、ネガティブなど、何らかの指標によって分類する研究が多く行われている(Pang 2005)。また、読み手にとって有益な情報として helpfulness を推測する研究がある(S. Kim et al. 2006)。機械学習による自動分類器が利用され、その素性として書き手が付与したレビュースコア、レビューの長さ、N-gram 等が使われている。このように、従来の多くの研究では、レビューのポジティブ / ネガティブの対極性や helpfulness といった特定の側面に着目しており、レビュー内容に関する奥の深い分析はなされていない。

本研究は、主観的な文章の内容に関する分類と分析を行うことを目的としている点で、談話構造の分類の研究と関連している。例えば、チャット対話の対話構造を解析する目的で作成された談話タグが参考になる(石崎 1998; 徳永 2005)。また、大野ら(2005)はネットオークションの出品者評価コメントの自動要約の事前調査として、18 人の出品者の 10 コメントを取り出し、出品者評価コメントの内容を調査している。出品者の「対応」とやりとりされた「商品」について連絡、謝罪、到着など大きく 13 種類のトピックに分類し、それぞれに属性値を持たせて情報を構造化している。本研究で我々が対象としているのは、サービスやトラベルを含めた多種多様な商品に対するレビューであり、この分類はそのまま当てはまらない。また、対象商品が幅広いが故に我々の分類項目は大野らの分類に比べて汎用的な枠組みとなっている。

## 3. レビューには何が書かれているか

レビューには、商品の感想のみならず、買い手の購買背景、商品の描写、売り手への要望、将来の買い手へのメッセージなど多種多様な事柄が書かれている。これまでのレビュー分析の研究では商品の感想を対象としたものが中心となっているが、レビューを有効活用するために、まずはレビューには何が書かれているのかをしっかりと見つめ直すことが重要だと考える。

多種多様な商品のレビューに共通の分類を構築するために、日本国内最大規模のインターネットショッピングサイトである楽天市場の 15 カテゴリーの商品レビューから各 100 文、国内最大規模の宿泊施設の予約サイトである楽天トラベルのレビューから 500 文をランダムにサンプリングし、分析対象とした。分析対象を文単位としたのは、レビュー全体を対象にすると多様な事柄が包含され分析が複雑になるためである。楽天市場の分析対象は、商品カテゴリーのトップジャンルから最低 1 カテゴリーとなるように、商品数を参考にして選んだ(表 1)。まず、分析者(本論文著者 2 名)が分析対象の文を読み、内容に関する分類項目を作成した。次に、新たな作業者を 1 名加え 3 名で分類項目に対して文を分類した。この段階でも分類項目の修正などは行った。作成した分類項目を次頁の表 2 に示す。

表 1: 調査した楽天市場の商品種類

トップジャンル	商品種類
ファッション	バッグ、レインコート
キッズ・ベビー・玩具	ぬいぐるみ
インテリア・日用雑貨	チェスト、ボールペン
食品・飲料・お酒	ワイン、チーズケーキ
本・音楽・ゲーム	邦楽 (ロック・ポップス)
美容・健康	香水、薬
スポーツ・アウトドア	自転車
家電・パソコン・通信	ノートパソコン、掃除機
ペット・花・DIY	種子
車・バイク・サービス	保険

### 3. 1. レビュー内容の分類項目

インターネットショッピングにおいて重要な役割を担うのは、商品、売り手、買い手の3つである。実際のレビューでも、商品の種類によらず、ほとんどのレビューの内容がこの3項目に関連した内容であった。本研究では、商品、売り手、買い手を以下のように定義する。商品とは「売するために・作られた、陳列された物であり、具象物ではないサービス・情報や、金融関係のものをも指す」<sup>1)</sup>。売り手とは「商品を販売している店舗、商品を生み出した企業、ショッピングサイト運営者(本データの場合は楽天(株)及び楽天トラベル(株))」とする。買い手とは「商品を購入した人物とそれを手にした人物」とする。商品を手にした人物とは、例えば、買い手が商品を友人にプレゼントした時の、友人がそれに当たる。この3つの大項目をレビュー内容の分析の大きな軸とし、さらに詳細な分類項目を作成した。分類項目の詳細な説明に関しては(安藤 2013)を参照された。

### 3. 2 レビュー文の分類

表2に示したように、レビュー内容について、3つの大項目、23の詳細項目からなる分類を作成した。これらに対し、先に挙げた2,000文を分類する作業を3名の分析者が分担して行った。実際には、1文には様々なことが記述されているため、各文を詳細項目に分けるのではなく、文中の対応する部分に対して表2の詳細項目をタグ付けするという方法で分析を行った。従って、1文に複数のタグが付くことがある。このタグ付けの精度を測るために宿泊施設のレビュー100文の正解タグデータを作成し、各分析者のタグとの一致度を調べた。範囲と分類の詳細項目の両方が一致した時のみを正解とした場合の一致度が69~92%であった。典型的な判定のずれや細かい範囲の違いなども含まれており、23種類の項目への分類としては高い一致度だと考えている。また、サンプル文のうち、「未分類」、「判定不可能」は商品カテゴリー毎に0~3.4%で、汎用性が高いことがわかる。

## 4. インパクトとの関係性

各詳細項目が、どのくらい読み手の購買意思決定に影響を及ぼしているかを調べたい。前節の分類作業を拡張し、楽天トラベルのレビューからランダムサンプリングした3,000文を新たに分類すると共に、読み手によるインパクト判定を行った。インパクト判定とは、「そのホテルに行きたい、または、行きたくないと思わせる文」を「(ポジティブ(P)/ネガティブ(N))なインパクトありの文」と定義したもので、実験協力者10名(男性、女性それぞれ5名)に判定を依頼した。本分析では10人中4人以上<sup>2)</sup>がインパクトありと判定した文を「(P/N)インパクト文」と定義し、以降の分析で使用する。

### 4. 1. 分類項目とインパクトの関係

分析対象の3,000文中で詳細項目の頻度が50以上の12項目について(P/N)インパクト文の割合を次頁の図1に示す。各項目の3,000文中における出現頻度は縦軸の括弧内の数値で示す。

<sup>1)</sup> (株)三省堂「新明解国語辞典 第四版」より

<sup>2)</sup> 4人以上がインパクトありと判定した文数は304文であり、この数が実験協力者のインパクトあり判定の平均文数に最も近かったため、4人を閾値とした。

(P/N)インパクト文の共通の現象として、商品に関する判定と絶対的事実、売り手に関する行動と判定、買い手に関するストーリーが5%以上含まれており、他の項目よりもよりインパクトに影響することがわかった。実例を見ると、商品に関する判定と絶対的事実、食事や部屋といったホテルに関する具体的なトピックの説明が多い。売り手に関する行動と判定は宿泊施設のレビューの場合、従業員に関するサービスの説明が大部分を占める。即ち、商品に関する(本データの場合は宿泊施設に関する)説明や印象がインパクト文に多く含まれる傾向が確認できた。一方、買い手の宿泊施設での行動、買い手が今後その宿

表2: レビューの分類項目 (Ver. 1)

大項目	詳細項目	説明
商品	行動意志	買い手が商品を手にした状態で、商品を使って今後どのような行動をする、というような意志表示
	行動	商品を含んだ買い手の行動
	結果	購入商品が完全に売り手の手を離れ、買い手の手で使用した時の行動・状態
	判定	商品を使用・購入した時の買い手の気持ち。買い手の感情を通した商品の判定(快・不快など)。商品を使用・購入した時の商品に対する判定
	未判定	商品を使用する前で評価していないもの
	推察	商品に対する経験を越えた推察や一般論
	スコア	星の判定に関する記述
	絶対的事実	有無、商品の特徴など、誰が見てもそうと認められるもの
	比較	同種もしくは類似の商品に関するコメント
	関連情報	商品に近い別商品の情報。同じメーカーの別商品比較以外の、商品以外の記述
売り手	行動	商品を売却する時の売り手の具体的な行動
	やりとり	売り手と買い手の言動のやりとりのうち、売り手の具体的な行動を除いたもの
	結果	商品を使用・購入した時の売り手の行動の結果
	判定	商品を使用・購入した時の売り手の行動やサイトに対する感想。買い手の感情を通した売り手の判定
	購入意志	店舗に対して、今後どのような行動をします、というような意志表示
	応援・御礼	具体的な記述のないポジティブな表現。応援メッセージ・御礼など。
	暴言	具体的な記述がないネガティブな表現。暴言・罵りなど。
買い手	要望	店舗・メーカーなどへの要望・苦情。商品への物足りなさ、良さの記述
	購入理由	購入決定した理由。買い手の利用目的(場面)。購入後の用途
	購入条件	購入の条件。プランなども含む
	予測・期待	商品を実際に手にしていない時のイメージ
	回数	商品購入や店舗の利用回数
ストーリー	商品を買う前の状態の説明。日ごろの習慣	

泊施設を利用するか否か、宿泊施設に対する謝意といった、買い手自身の宿泊施設に対する印象・行動の一部は、読み手にとっては重要ではない傾向が確認できた。売り手への要望はNインパクト文に対してのみ、買い手の回数はPインパクト文に対してのみ影響を与える要因であった。

#### 4. 2. 分類、トピック、インパクトの関係

前節では、詳細項目毎にインパクトとの関係を分析したが、その結果は食事や部屋などのトピックにも関係することがわかった。本節では、インパクト文の特徴を詳細に分析すべく、トピックを加えてインパクトとの関連性について考察する。トピックは楽天トラベルで定義されている6種類(立地、部屋、食事、風呂、サービス、設備・アメニティ)を利用し、詳細項目とトピックの組の中で(P/N)インパクト文に出現した頻度が10以上となったものを分析対象とする。詳細項目とトピックが同時に出現した頻度分布に独立性を仮定すると、それぞれの組の期待頻度が計算できる。期待頻度に対し120%以上の出現があった組、もしくは80%以下の頻度しかなかった組を表3に示す。

商品に関する絶対的事実は、部屋、風呂、サービスのNインパクト文に多く出現する傾向がわかった。部屋と風呂に関する文には「(ゴミ/髪の毛)が落ちている」といった不衛生さの記述が多く見られた。サービスのNインパクト文には「フロントにスタッフがいない」といった絶対的事実が多く存在する一方で、Pインパクト文には絶対的事実が含まれない傾向が確認できた。それよりも「仲居さんの対応が丁寧だ」といった宿泊施設のスタッフの行動や評価の記述が多いことが観察できた。

売り手に関する判定では、食事のPインパクト文が多く出現する傾向がわかった。食事の場面ではスタッフが給仕することが多く、気持ちの良い食事のためには適切なスタッフの行動が重視される傾向が観察できた。また、売り手に関する行動では、サービスのNインパクト文の頻度が多く見られた。「納得のいく説明がない」などのスタッフの行動が、その宿泊施設に行きたくなくなるような強いインパクトを与えることがわかった。

買い手に関する行動ではサービスのNインパクト文の頻度が多く、「スタッフに話しかけづらい」といった何かができなかった内容や、「早く到着した」際の対応が不満といった、スタッフの評価と共に出現する例が見られた。

### 5. インパクトに影響を与える要因

前節までの分析ではレビュー内容の分類、トピックといった概念的な分類に焦点を当ててきた。本節では、文内容に焦点を移し、ポジティブな内容の文に対して読み手にインパクトを与える要因について分析する。分析対象は前節と同じく楽天トラベルのレビューからランダムサンプリングをした3,000文である。Pインパクト文では、例えば、「朝食が韓国料理でした」といった商品の特殊性や、「夕食はキムチ、チゲ、チジミなどの韓国料理が盛りだくさんでした」といった描写の具体性、「心から堪能しました」といった買い手の満足度合いが読み手のインパクト判定に影響を与えることが経験的にわかった。本節ではこれらがインパクトの判定に与える影響を定量的に分析する。

#### 5. 1. 調査の設定

商品の特殊性、描写の具体性、買い手の満足度合いがインパクト判定に与える影響を判定数が比較的多く含まれる食事、風

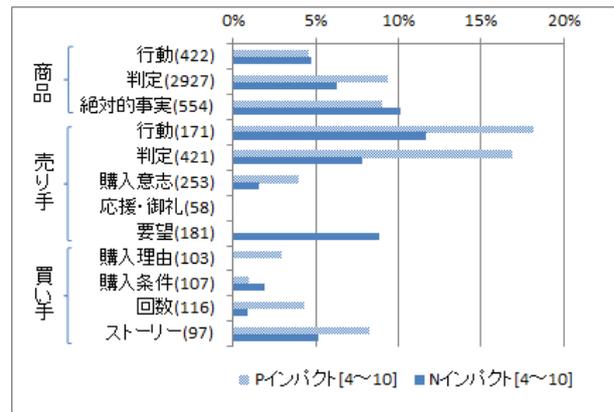


図1: 項目がP/Nインパクト文に含まれる割合

表3: インパクト文に多く含まれる分類&トピック

項目	トピック	項目&トピックの頻度	P/N	イバ外文での頻度(期待頻度)	
商品	絶対的事実	部屋	97	N	22 (11)
		風呂	123	N	21 (15)
		サービス	109	N	16 (11)
		サービス	109	P	12 (21)
売り手	判定	サービス	402	P	92 (76)
	判定	食事	35	P	15 (7)
買い手	行動	サービス	157	N	20 (16)
買い手	行動	サービス	65	N	10 (6)

呂、サービスのトピックについて調査する。インパクト文の判定で実験協力者10人中インパクトありと判定した人数が多いほどより強いインパクトがあるものとし、その人数と上記に挙げた3つの要因の相対的順位の相関を調べる。これらの要因は絶対的に評価することが困難なため、20代から30代の5人の実験協力者に相対的判定を依頼した。Pインパクト判定を行った人数毎に(但し5人以上がインパクトありとした文は少ないため、5、6人の判定文、7~10人の判定文は各々ひとつのグループとする)サンプル文を1文ずつ抽出した7文からなるデータセットを10セット用意する。これらの文に対して商品の特殊性、描写の具体性、書き手の満足度合いについて1位から7位まで相対的なランク付けを依頼した(但し同ランクを許す)。

#### 5. 2. 結果

実験協力者5名が付与したランクの平均とインパクト判定数で順位付けした結果について、トピックごとに相関係数を求めた(図2-4)。すべてのトピック、要因において相関係数が0.8以上となり、相関性が高い結果となった。即ち、商品の特殊性、描写の具体性、書き手の満足度合いがインパクトに影響を与えていることが確認できた。各トピックで最もインパクト文との相関性が高かったのは、食事とサービスが書き手の満足度合いで、風呂が商品の特殊性であった。

実際の文を分析した結果、以下のことが分かった。食事に関するレビュー文には「おいしい」や「最高に美味しい」といった書き手の満足度合いの強弱が分かるような表現が多く含まれ、その強弱がインパクトに影響を与えていると考えられる。食事

の特殊性には、「手作り感、地元感が表に出ている最高のバイキング」といったインパクトに影響を与えるものと、「韓国料理がある」といったそれほど強いインパクトを与えないものが見られたため、満足度合いよりも相関性への影響が少なかったものと考えられる。サービスは、宿泊施設のスタッフに対する謝意が多く含まれており、食事と同様書き手の満足度合いの強弱が読み手の判断に影響していると考えられる。サービスの特殊性には、「サプライズ」、「幼児用の草履の用意」や「チェックイン時のお茶のサービス」など幅広いもの含まれているため、特殊性のランク付けに人によって違いが現れた。風呂はリフレッシュや癒しの場であり、「疲れがとれる」、「気持ちが良い」といった書き手の満足度にはあまり強弱が見られなかった。それよりも「露天ジャグジー」、「檜風呂」といった特殊性の方が読み手にとって価値ある情報となることわかった。

これらの結果からトピックによってインパクトの要因が異なる可能性が確認できた。将来的にインパクト文を自動抽出する際には、トピック毎にその要因を捉えた上で処理すれば、読み手の購買意欲を高めるような情報の抽出が期待できよう。

## 6. まとめ

レビューの分析に関する従来の研究では、ポジティブ / ネガティブなど、レビューの特定の側面に着目した研究が主であった。本研究では、レビューに何が書かれているのかに着目し、23項目の分類を作成した。これを基にレビューに対して分類作業を行ったところ、幅広いジャンルの商品に対して95%以上の汎用性があることがわかった。また、楽天トラベルのレビュー3,000文に対して分類した結果と読み手の購買意欲をかき立てるようなインパクトとの関係を調査し、商品に関する絶対的事実がポジティブ、ネガティブのインパクトに影響を与えやすいことを示した。さらに、インパクトに影響を与える要因として、商品の特殊性、描写の具体性、書き手の満足度合いを示し、これらがインパクト判定に与える影響を定量的に示した。

本論文ではインパクトの調査をしたテキストが楽天トラベルに限られていたため、今後はテキストを他の商品レビューにも拡大し、インパクトに影響を与える要因が他の商品に関するテキストにも有効であることを調査したいと考えている。

## 謝辞

本研究を進めるにあたり、データを提供してくださいました楽天技術研究所に心より御礼申し上げます。

## 参考文献

- 安藤まや; 関根聡, レビューには何が書かれているのか?, ALAGIN & NLP 若手の会合同シンポジウム, <[http://yans.anlp.jp/include/netcommons\\_file.php?path=/announcement/1/NLP\\_yans\\_2013\\_ando.pdf](http://yans.anlp.jp/include/netcommons_file.php?path=/announcement/1/NLP_yans_2013_ando.pdf)>(2014/1/24 アクセス), 2013.
- 石崎雅人; 加藤回昭, 多人数対話の特徴分析, 言語処理学会第4回年次大会発表論文集, pp.402-405, 1998.
- 大野華子; 楠村幸貴; 土方嘉徳; 西田正吾, 社会的関係を用いたネットオークションの評価コメントの自動要約, 電子情報通信学会論文誌 D-I, J88-D-I No.3, pp.668-683, 2005.

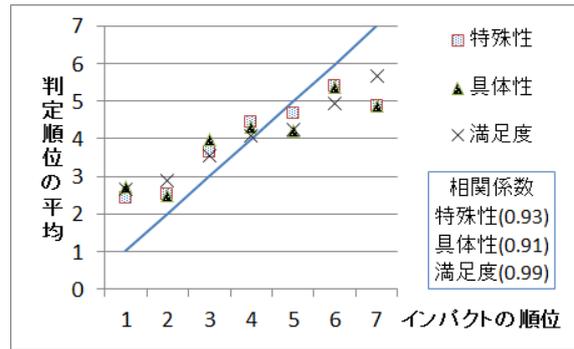


図2：食事の相関

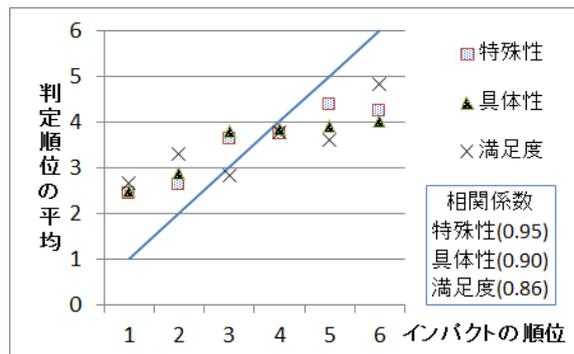


図3：風呂の相関

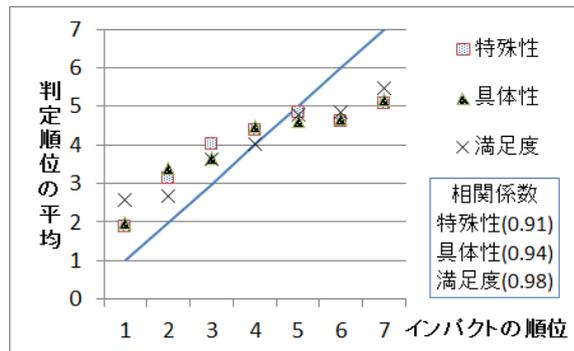


図4：サービスの相関

徳永泰浩・乾健太郎・松本裕治, チャット対話における発言間の継続関係と応答関係の同定, 自然言語処理 12(1), pp.79-105, 2005.

Bo Pang and Lillian Lee, "Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales", Proceedings of the Association for Computational Linguistics, 2005.

M. Koppel and J. Schler, "The Importance of Neutral Examples in Learning Sentiment", Computational Intelligence, 22(2), 2006.

S.-M. Kim, P. Pantel, T. Chklovski and M. Pennacchiotti, "Automatically assessing review helpfulness", Proceedings of the Conference on EMNLP, 2006.